

WHAT IS CLAIMED IS:

1 1. A method for identifying oligonucleotide sequences suitable for the
2 amplification of a unique sequence within a genomic region of interest, said method
3 comprising the steps of:
4 executing a first process on a digital computer to identify repeat sequences
5 that occur within said genomic region of interest;
6 executing a second process on a digital computer to compare repeat
7 sequence-free subsequences within said genomic region of interest to a nucleotide sequence
8 database, whereby nucleotide sequences within said nucleotide sequence database that are
9 substantially similar to said repeat sequence-free subsequences are identified;
10 executing a third process on a digital computer to identify oligonucleotide
11 sequences that are suitable for use as primers in an amplification reaction to amplify a
12 product within any of said repeat sequence-free subsequences for which a defined number of
13 substantially similar sequences are identified in said nucleotide sequence database; and
14 outputting said oligonucleotide sequences.

1 2. The method of claim 1, wherein said genomic region is from a human
2 genome.

1 3. The method of claim 1, wherein said number of substantially similar
2 sequences is zero.

1 4. The method of claim 1, wherein said oligonucleotide sequences are
2 outputted by displaying the sequences on a computer screen or on a computer printout.

1 5. The method of claim 1, wherein said oligonucleotide sequences are
2 outputted by executing a fourth process on a digital computer to direct the synthesis of
3 oligonucleotide primers comprising said oligonucleotide sequences.

1 6. The method of claim 5, wherein said computer directs the synthesis of
2 said oligonucleotide primers by ordering said synthesis from an external source.

1 7. The method of claim 5, wherein said computer is in communication
2 with an oligonucleotide synthesizer, and wherein said computer directs the synthesis of said
3 oligonucleotide primers by said synthesizer.

1 8. The method of claim 1, wherein said substantially similar sequences
2 are at least about 50% identical to said repeat sequence-free subsequences.

1 9. The method of claim 1, wherein said substantially similar sequences
2 are at least about 70% identical to said repeat sequence-free subsequences.

1 10. The method of claim 1, wherein said substantially similar sequences
2 are at least about 90% identical to said repeat sequence-free subsequences.

1 11. The method of claim 1, wherein said first process is executed using
2 Repeat Masker software.

1 12. The method of claim 1, wherein said second process is executed using
2 a BLAST algorithm.

1 13. The method of claim 1, wherein said third process is executed using
2 Primer3 software.

1 14. The method of claim 5, further comprising producing an amplification
2 product using said oligonucleotide primers.

1 15. The method of claim 14, wherein said amplification product is a FISH
2 probe.

1 16. The method of claim 15, wherein said FISH probe is fluorescently
2 labeled.

1 17. The method of claim 14, wherein said amplification product is an array
2 CGH target.

1 18. A method for identifying oligonucleotide sequences suitable for the
2 amplification of a unique sequence within a genomic region of interest, said method
3 comprising the steps of:

4 analyzing a genomic nucleotide sequence that encompasses said genomic
5 region of interest to identify repeat sequences within said genomic region;

6 comparing at least one repeat sequence-free subsequence within said genomic
7 nucleotide sequence to a nucleotide sequence database to identify sequences within said
8 database that are substantially similar to said repeat sequence-free subsequence;

9 for at least one of said repeat sequence-free subsequences for which a defined
10 number of substantially similar sequences are identified within said nucleotide sequence
11 database, selecting oligonucleotide sequences that are suitable for use as primers in an
12 amplification reaction to amplify a product within said repeat sequence-free subsequence.

1 19. The method of claim 18, wherein said genomic region is from a human
2 genome.

1 20. The method of claim 18, wherein said defined number of substantially
2 similar sequences is zero.

1 21. The method of claim 18, further comprising displaying said
2 oligonucleotide sequences on a computer screen or on a computer printout.

1 22. The method of claim 18, further comprising directing the synthesis of
2 oligonucleotide primers comprising said oligonucleotide sequences.

1 23. The method of claim 22, wherein said synthesis is directed by ordering
2 the synthesis of said primers from an external source.

1 24. The method of claim 18, wherein said substantially similar sequences
2 are at least about 50% identical to said repeat sequence-free subsequences.

1 25. The method of claim 18, wherein said substantially similar sequences
2 are at least about 70% identical to said repeat sequence-free subsequences.

1 26. The method of claim 18, wherein said substantially similar sequences
2 are at least about 90% identical to said repeat sequence-free subsequences.

1 27. The method of claim 18, wherein the identification of repeat sequences
2 within said genomic region is performed using Repeat Masker software.

1 28. The method of claim 18, wherein the comparison of said at least one
2 repeat sequence-free subsequence with said genome database is performed using a BLAST
3 algorithm.

1 29. The method of claim 18, wherein said oligonucleotide sequences are
2 selected using Primer3 software.

1 30. The method of claim 22, further comprising generating an
2 amplification product using said oligonucleotide primers.

1 31. The method of claim 30, wherein said amplification product is a FISH
2 probe.

1 32. The method of claim 31, wherein said FISH probe is fluorescently
2 labeled.

1 33. The method of claim 30, wherein said amplification product is an array
2 CGH target.

1 34. A computer program product designing and outputting oligonucleotide
2 sequences suitable for use as primers to amplify unique sequences within a genomic region
3 of interest, said computer program product comprising:
4 a storage structure having computer program code embodied therein, said
5 computer program code comprising:
6 computer program code for causing a computer to analyze a nucleotide
7 sequence encompassing said genomic region of interest to identify repeat sequences within
8 said nucleotide sequence;
9 computer program code for causing a computer to, for each subsequence of
10 said nucleotide sequence that does not contain any of said repeat sequences, compare said
11 subsequence against a nucleotide sequence database to identify nucleotide sequences within
12 said database that are substantially similar to said subsequence;
13 computer program code for causing a computer to, for each of said
14 subsequences for which a defined number of substantially similar sequences are found in said
15 database, identify oligonucleotide sequences suitable for use as primers in an amplification
16 reaction to amplify a product within said subsequence; and
17 computer program code for outputting said oligonucleotide sequences.

1 35. The method of claim 34, wherein said defined number of substantially
2 similar sequences is zero.

1 36. The method of claim 34, wherein said substantially similar sequences
2 are at least about 50% identical to said subsequences.

1 37. The method of claim 34, wherein said substantially similar sequences
2 are at least about 70% identical to said subsequences.

1 38. The method of claim 34, wherein said substantially similar sequences
2 are at least about 90% identical to said subsequences.

1 39. A method for identifying genes within a genomic region of interest,
2 said method comprising the steps of:
3 executing a first process on a digital computer to identify repeat sequences
4 that occur within said genomic region of interest;
5 executing a second process on a digital computer to compare repeat sequence-
6 free subsequences within said genomic region of interest to a nucleotide sequence database,
7 whereby nucleotide sequences within said nucleotide sequence database that are substantially
8 similar to said repeat sequence-free subsequences are identified;
9 executing a third process on a digital computer to select repeat sequence-free
10 subsequences having no substantially similar sequences to identify a repeat sequence-free
11 subsequence may represent a gene family.
12 identify oligonucleotide sequences that are suitable for use as primers in an
13 amplification reaction to amplify a product within any of said repeat sequence-free
14 subsequences for which a defined number of substantially similar sequences are identified in
15 said nucleotide sequence database; and
16 outputting said oligonucleotide sequences.